Obtaining a Snellius Account

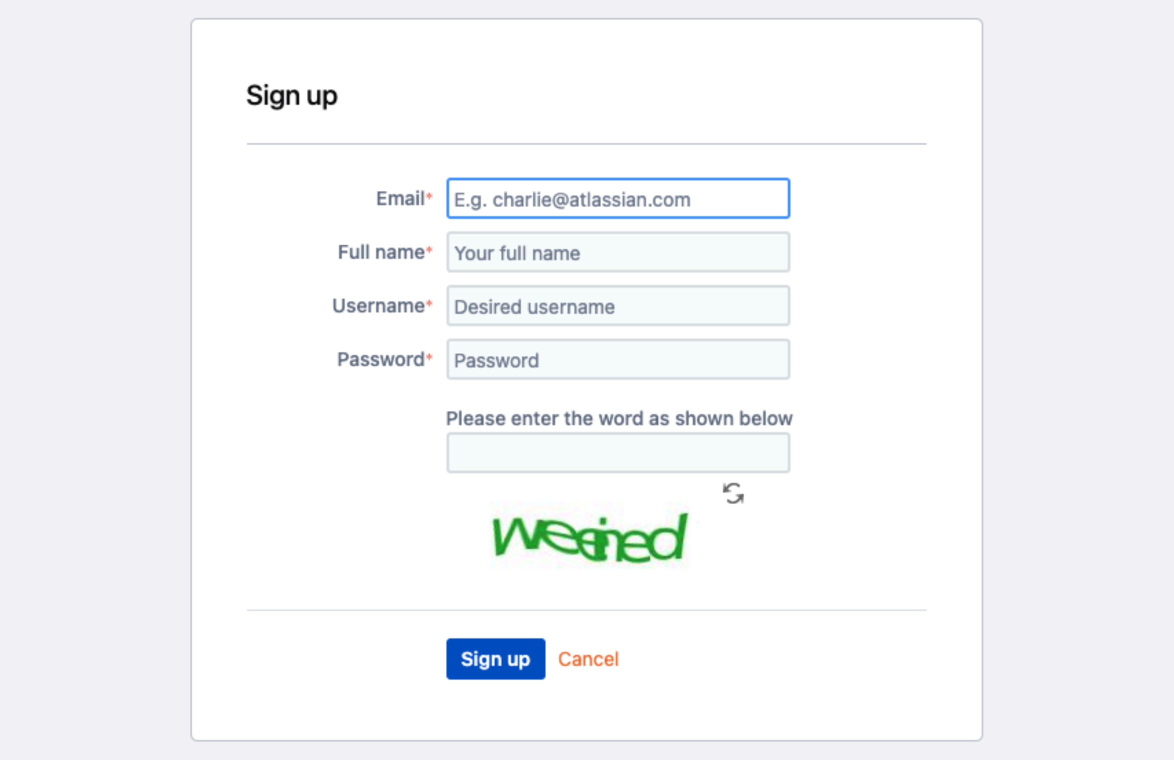
For PGC Users

Sarah Colbert, B.A., sarah.colbert@icahn.mssm.edu
Prof. Stephan Ripke, M.D., Ph.D., Group Leader, sripke@broadinstitute.org

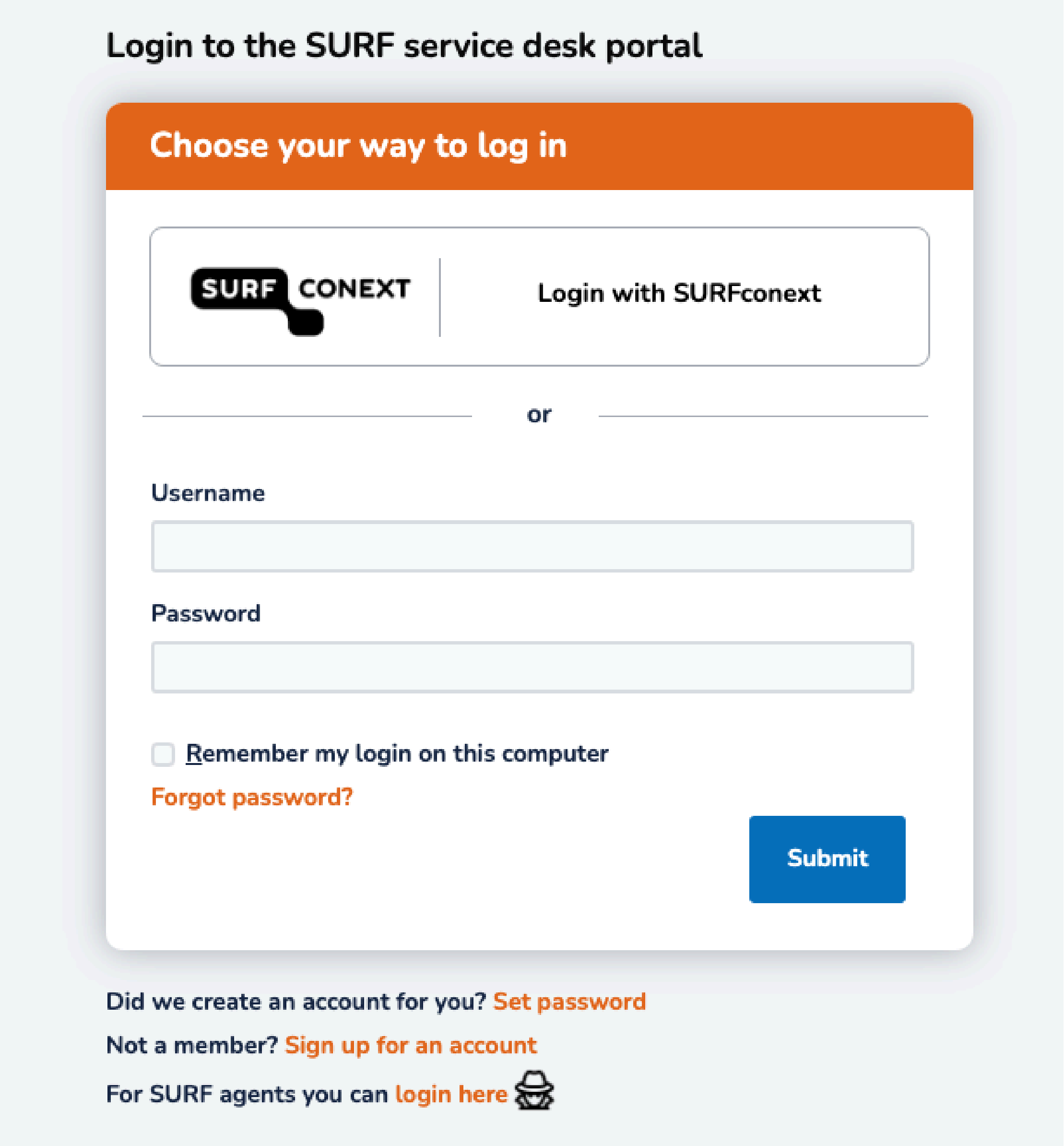
PRE-STEPS: Prior to applying for an account, please make sure that you meet all of the following requirements!

- ☑ I have project approval/confirmation email from my WG chair OR a Phase 2 approval email.
- ☑ I have completed the required Snellius user tutorials and I understand how to run jobs efficiently on the Snellius cluster.
  - ☑ Connecting to Snellius tutorial
  - ☑ Snellius batch system tutorial
  - ☑ Snellius efficient CPU jobs tutorial
- ☑ I understand that Snellius must be acknowledged in publications using the following text: “We thank SURF (www.surf.nl) and VU University Amsterdam (https://vu.nl/en) for the support in using the National Supercomputer Snellius.”

STEP 1: If you do not already have a Snellius account, you will need to apply for one. To do so, please sign up for Jira so that you can login to the service desk. The signup page should look like:



STEP 2: After creating an account, you can use this link to login to the SURFsara Service Desk which should look like:



Login to the SURF service desk portal

Choose your way to log in

Login with SURFconext

or

Username

Password

☐ Remember my login on this computer

Forgot password?

Submit

Did we create an account for you? Set password

Not a member? Sign up for an account

For SURF agents you can login here

STEP 3: Upon logging in via the link above, you will be directed to the page used to submit your request for an account. From the "Choose your organization" dropdown menu please select "Psychiatric genomics consortium (PGC/GCC)" and in the summary write "Snellius account request". If your request is for the purpose of uploading your own data to the PGC Data Receiving Committee, please select "Snellius CPU" for the type of request and indicate this in the Description field, along with the working group(s) involved.

In case you don't have a very clear estimate please request 5,000 SBUs[1] for the start. You may also find this document helpful in estimating SBUs.

Important:

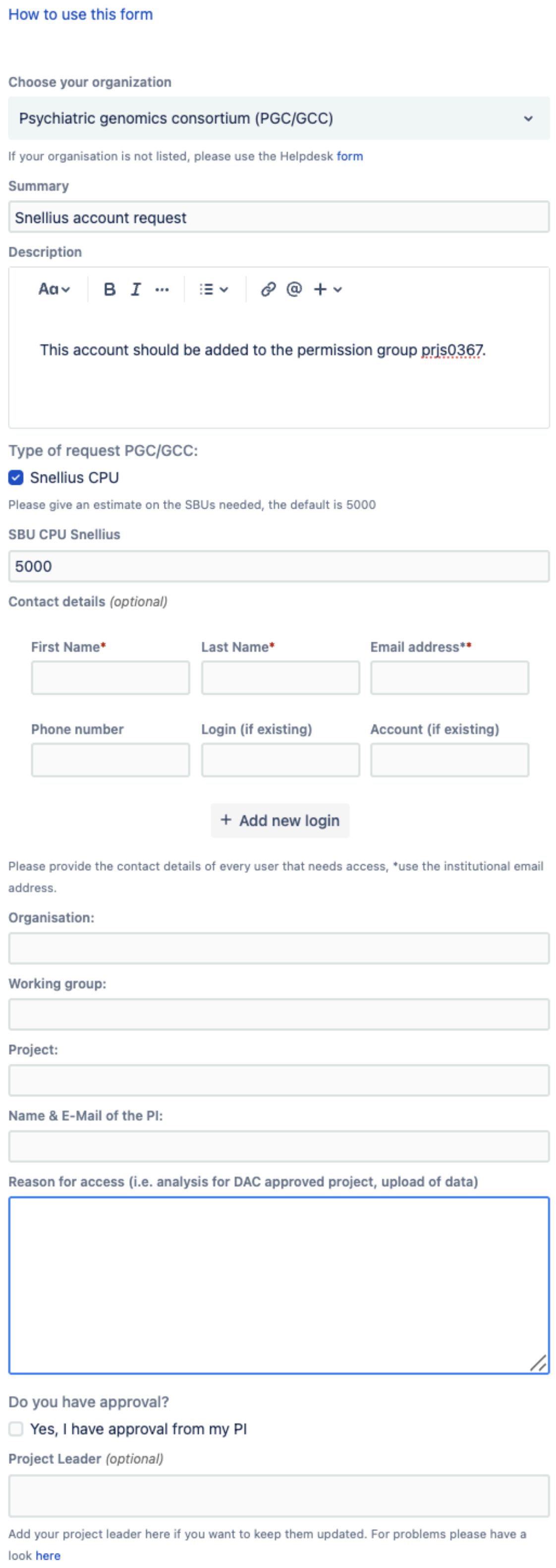
1) Please also name the project (e.g. DAC approved project), working group and your PI.

---

[1] While unlikely, if you suspect you may have exceeded your budget (e.g. SLURM warns you about low SBUs), you can check using the "budget-overview" command. By default, extra SBUs will be added in increments of 5 or 10K. However, if you need a higher amount, you can submit a new request using the same form. More information on accounting is available here.

2) Please add as the first line of the Description that the user should be added to permission group prjs0367

3) Please specify your reason for access (i.e. to perform analysis for a DAC approved project or to upload data). Note that you will need to upload a proof of this later (see earlier comment).



How to use this form

Choose your organization

Psychiatric genomics consortium (PGC/GCC)

If your organisation is not listed, please use the Helpdesk form

Summary

Snellius account request

Description

This account should be added to the permission group prjs0367.

Type of request PGC/GCC:

☑ Snellius CPU

Please give an estimate on the SBUs needed, the default is 5000

SBU CPU Snellius

5000

Contact details *(optional)*

First Name*

Last Name*

Email address**

Phone number

Login (if existing)

Account (if existing)

Please provide the contact details of every user that needs access, *use the institutional email address.

Organisation:

Working group:

Project:

Name & E-Mail of the PI:

Reason for access (i.e. analysis for DAC approved project, upload of data)

Do you have approval?

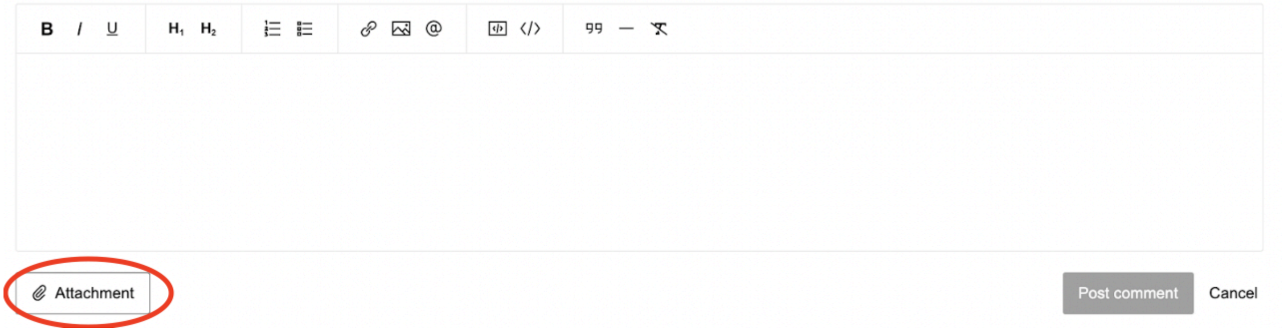
☐ Yes, I have approval from my PI

Project Leader *(optional)*

Add your project leader here if you want to keep them updated. For problems please have a look here

You should receive a confirmation email that a ticket has been created shortly after submitting the request. If you do not receive a reply within 48 hours, don't hesitate to email Danielle Posthuma (danielle.posthuma@vu.nl) and Walter Pirovano (w.a.pirovano@vu.nl).

Once the ticket has been created, it is necessary to provide the project approval/confirmation email from your WG chair (in case you need to perform analysis for a DAC approved project) OR a Phase 2 approval email (subject line "Data Inquiry and Intake Forms Approved") (in case you need to upload data). You can do this by uploading the approval in the ticket (using the Attachment button, see below) or by sending the approval as an email attachment (in reply to the ticket confirmation email by SURF).



NOTE: In case you need would like to request additional SBUs (because you have used all initially provided SBUs, you can request additional SBUs using the same (account application) webform. However please mention in the description that you need additional SBUs (and not a new account).

STEP 4: Once you have an account, you can login via command line to the Snellius server. For the purpose of this documentation we always assume your account to be `user` and for the '$' at the beginning of all code snippets / commands to be your prompt:

```
$ ssh user@doornode.surfsara.nl
```

Please note that the "doornode" host is not suitable for data upload/download using protocols such as SFTP, SCP, or Rsync (i.e. this is needed for uploading data!!). To have full access with SCP/SFTP and other protocols, you will need to provide your public IP address (you can list many, e.g. if you are using different workspaces or VPNs) to servicedesk@surf.nl. You can determine your public IP address by running the following command or by visiting the following URL

```
$ curl http://ipecho.net/plain; echo
```

Once your IP has been cleared and granted access, you can use the host `"snellius.surf.nl"` to log in with full access. To do so, you can use the following command:

```
$ ssh user@snellius.surf.nl
```

STEP 5: you might consider applying for a project space (your homedir space is limited). But only one per login possible:

https://pgcdataaccess.formstack.com/forms/pgcdac_partition_request

Please let us know if you have questions or run into any issues:
sripke@broadinstitute.org
sarah.colbert@icahn.mssm.edu

A note on file size limitations (during file uploads)

Some users have experienced issues in uploading files with OOD given due to an imposed file size limitation of 10 GB. As a solution to overcome this, you can use programs like FileZilla, WinSCP or Cyberduck (Mac). These do not have any file limit imposed (AFAIK).

However, **Rsync** is way more reliable and allows restarting a failed download even in the middle of a partially uploaded file. In that case, use the following command to upload data from a terminal on your computer.

```
rsync -a SOURCE user@snellius.surf.nl:/home/user/TARGET
```

SOURCE should be the path to the file/s that you want to upload. It accepts the asterisk (and full regex also). Thus, if i.e. you want to transfer all the zip files in folder *`/home/youruser/data`*, you could write:

```
rsync -a /home/youruser/data/*.zip
user@snellius.surf.nl:/home/user/newfolder
```

CAUTION: Rsync discriminates between paths ending in a slash and paths without it.

```
#1      rsync -a /home/youruser/data snellius.surf.nl:/home/user

#2      rsync -a /home/youruser/data snellius.surf.nl:/home/user
```

The first option, without a slash, will copy the content of the folder "data" directly into your home directory. The second option will create a folder called "data" in your home directory with all the files inside it.

From a terminal in OOD or on Snellius: In the case that your data is on a cluster or server, you need to turn the command around, thus placing the full SSH command user@host first:

```
#1      rsycn -a user@REMOTE:/path/on/the/remote /home/user/myfolder

#2      rsycn -a user@REMOTE:/path/on/the/remote myfolder

#3      rsycn -a user@REMOTE:/path/on/the/remote $(pwd)
```

The first version would place the content of the remote path *`/path/on/the/remote`* into your folder *`myfolder`* on Snellius. Both versions are identical. I like to add the full path as it can be used in scripts and does not depend on where you are when you run the command. The second assumes that you are in `/home/user`.